

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2003, 15:17:50 ; Search time 43 Seconds  
(without alignments)  
1693.016 Million cell updates/sec

Title: US-09-494-297-2  
Perfect score: 3945  
Sequence: 1 MKKTRFPNKLNTLNTQRYLS.....IAGISLGIWCIHTIRIRKHD 757

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	633.5	16.1	659	2 SA00043	adhesin - Streptococcus
2	603	15.3	638	2 S54418	fibronectin-binding
3	164	4.2	1039	2 T30856	protein F2 - Streptococcus
4	161	4.1	1386	2 AC1533	surface protein (L
5	159	4.0	1243	2 S60138	sex factor aggrega
6	159	4.0	1612	2 AB1347	probable peptidogl
7	157.5	4.0	5005	2 PB2884	hypothetical prote
8	156	4.0	1530	2 AH1396	peptidoglycan anch
9	155	3.9	832	2 AD1096	internalin protein
10	152.5	3.9	2013	2 AD1129	probable peptidogl
11	151	3.8	2367	2 S70172	toxin B - Clostrid
12	149	3.8	1315	2 T28679	fibronogen-binding
13	149	3.8	1946	2 AE1449	hypothetical prote
14	148.5	3.8	1883	2 G82875	hypothetical prote
15	147.5	3.7	2167	2 AF1489	cell wall-associat
16	145.5	3.7	1036	2 F90073	hypothetical prote
17	145	3.7	1036	2 T30311	S-layer protein -
18	144.5	3.7	940	2 AD1374	internalin protein
19	144.5	3.7	1092	2 T30214	fibronogen-binding
20	144.5	3.7	1578	2 AD1512	peptidoglycan bou
21	143	3.6	1185	2 A42404	collagen adhesin -
22	143	3.6	1385	2 D89824	hypothetical prote
23	142.5	3.6	627	2 T03821	prolipoprotein p65
24	141.5	3.6	1622	2 AE1717	probable cell surf
25	141	3.6	3216	2 C90538	hypothetical prote
26	140.5	3.6	876	2 G90592	hypothetical prote
27	139.5	3.5	888	2 E82885	hypothetical prote
28	139.5	3.5	1167	2 F71909	hypothetical prote
29	139	3.5	621	2 B64546	chaperone and heat

30	139	3.5	1433	1 A36734	bailliopeptidase F
31	139	3.5	2366	2 S10317	toxin B - Clostrid
32	138.5	3.5	625	2 T03837	surface lipoprotei
33	138.5	3.5	1582	2 AC1153	adhesin homolog 1m
34	138.5	3.5	1983	2 G86643	hypothetical prote
35	138.5	3.5	2013	2 A11489	probable peptidogl
36	138.5	3.5	2484	2 T26215	hypothetical prote
37	138.5	3.5	2607	2 T26215	hypothetical prote
38	137	3.5	1177	2 I64233	hypothetical prote
39	135.5	3.4	3283	2 AC1018	large repetitive p
40	135	3.4	926	2 D86897	hypothetical prote
41	135	3.4	1117	2 S33851	fibronectin-bindin
42	135	3.4	1176	2 A33856	surface-layer 125k
43	135	3.4	1228	2 I40468	surface layer prot
44	135	3.4	1272	2 C90593	hypothetical prote
45	135	3.4	1301	2 S51323	SAC3 protein - yea

## ALIGNMENTS

## RESULT 1

S40043  
adhesin - Streptococcus pyogenes  
C:Species: Streptococcus pyogenes  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 28-May-1999  
C:Accession: S40043  
R:Seila, S.; Avty, A.; Tovli, A.; Burslein, I.; Caparon, G.; Hanski, E.  
Mol. Microbiol. 10, 1049-1055, 1993  
A:Title: Protein F: an adhesin of Streptococcus pyogenes binds fibronectin via two di  
A:Reference number: S40043; MUID:95020565; PMID:7934855  
A:Accession: S40043  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-659 <SEL>  
A:Cross-references: GB:L10919; NID:9425479; PIDN:AAA26964.1; PID:9425480

## Query Match

Best Local Similarity 26.1%; Score 633.5; DB 2; Length 659;

Matches 203; Conservative 128; Mismatches 237; Indels 211; Gaps 30;

QY	20	SKNSRFFVTTLGVLMIFALVTSMVGAKTVEGLVESSTPNALINDSSSEYRMYESYV	79
DB	24	TKRRRFAVTLGVFMFLACAGALGPGQVAAAEKTVPSHSSP--NPEFMYGIDAYG	81
QY	80	RGHPRYKQFRAHDLRVNLEGSRSYQVYCFNLKAPPLGSDSSVKKWKYRKHDGISTKRED	139
DB	82	KEYPGYNIMTRYHDLRVNLSRSYQVYCFNLSNVYPSQKNSFIKMFKEIGNGKSFVD	141
QY	140	YAMSPRITGDELNOKLRAVMYNGHPQNNNGIMEGLEPNAIRVQGEAVWYVSDNAPISNP	199
DB	142	YAHYTKLKELEQRLSLLYNAYPNDANGYKKGLEHNAITVTOYAAWHSDNSOYGF-	200
QY	200	DESFRESESNLVSTQSLSIMROALKQIDLPMLAKRMPQVDDQLSTFSESDGDKYN	259
DB	201	ETIMSESEKKEGKISSQYTLNMEALKLIDPLELTAANKIPSGRLNIFESE-----N	254
QY	260	KGYONLSGLVPTKPPPGDPMPNPQPTTSVILIRKAYAGYSKLLLEGATL----	315
DB	265	EAYONLLSAEYVDPDPKPGE--TSEHNPKTPE-----LDGPIPDPKHP	298
QY	316	GDNVNSPFAVRESSDIERIELSDGYTLTLELNSPAGYIAEPIITFVEAGKYTTIDG	375
DB	299	DNLEPTLPVYMLD---GEVY-----DEVPSESLPALPLMP-----LDG	337
QY	376	KQIENPKIVE-PYSVEAYNDFEEFSVLTQNYAKFYAKNKNSSQVYCFNADLSP	434
DB	338	QEV--PEKPSIDPLEVPRYE-----FNKKDQSP	364
QY	435	PDSEDGKTYMPDFTTGKVKY--THIAGDLKRYIVKPRDTPDFLKHKKVIEKGRK	493
DB	365	LAGE-----SGEYETITEVYGNG-----QNPVDIDK-----KLPNETGF--S	399



```

OY 245 ---OLSFSEDEKDKXNK-----GYONLLSG-----GLVTPKPT 277
D 311 LOGAFVVKYTKGRADQSGKELIVOSHLSFRNMGSTYTLRPNHVSFTNELALSPSGSG 370
OY 278 PGDPMPNPOTTSVLRKAYIGDYKSL-----LEGATLQTLGDNVNSFOARFSSND 331
D 371 SCGSEF--TKPSITVANLKRVAQLRFKFKSTDNVPLPEAFAELRSSNGNS-QKLEASNT 427
OY 332 IGE--RIELSDGYITLTELNSPAGYSTA-----EPIIF-KVEA 366
D 428 QGELHFKDLTSGYIDLKETAPKGYQVTEKLATVTDVTKRPAEOMKWEKPHSFVVEA 487
OY 367 GKVTYIID-----KQOIEPNKKEIPEVPSYEA-----NDPEERSVLTQNYA 409
D 488 NKETIYNHKEITLFSKKKIMEDDRPDORAKIOVOLLONGQKHPNOIQE---YTKNDW 544
OY 410 KFYV-----AKNKGSSQVYVCFNADLSPDSEGGK-----TMPDPFTTGE 452
D 545 SYHFKDLPRKYDAKNOEKYKSV-----AEVKVP-----DGKVSYLGNIDFNTRETEFVEEQ 595
OY 453 VKTHHNGRDLFKYVPRDTPFLKHKKVLEKQ--YREKGALEYSGLETQRA- 509
D 596 NNENLEEGNAEIGQSGPKIIDEDTLTSPKSKKIMKNDTAENRPOAI-----QVQYAD 649
OY 510 --ATQALAYFTDSA---EIDKDKLDYHGFDMNDSTLAVAKILVEYADSNPQLTDL 564
D 650 GVAVEGQTRFISGSGNMSFEFKNLKYNSTG--NDIIVSKVEVTFPGD----- 698
OY 565 DFLIPNNKYQSLIGTQWHPEDLVDIRMEDKKEVIFVT--HNLTKRTV--TGLADRT 620
D 699 -----VYSSANDIIN-----TKREVITQCGPNLEIEETLPLESAGSGTT 738
OY 621 KDFHFEIE-----LNKKOELLSTQVTKDK--TNLEK-----GCK-----ATINLKH 661
D 739 -----TVEDSRPVDTLSGLSSEQSGDMTIEEDSAHIKFSKRDIDGKELAGTMBLRD 793
OY 662 GESLTLOG-----LPEGYSLVKEETDESGYK-----KYNSOEVAANATVSKT 703
D 794 SSGTITISWISDQVKNFYLMFGKYTF-VETAPDAVEIAATITFYVNEO--GGVTVNGK 850
OY 704 GITSDETLAEENKKEPVPTG---VDOKI 729
D 851 ATKGAHIVAVDAVKPKTSGOVVIDIEKL 880

RESULT 4
AC133
surface protein (LPXTG motif) [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AC133
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussaugel, O.; Entlian, R.D.; Fahl, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunsit, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Me
ok, C.; Schluter, T.; Stimes, N.; Tjertret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC133
A:Species: Listeria innocua
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1386 <GLAY>
A:Cross-references: GB:AL592022; PIDN:GAC96035.1; PID:916413254; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0803

```

Query Match 4.1% Score 161; DB 2; Length 1386;  
 Best Local Similarity 19.7%; Pred. No. 0.3;  
 Matches 163; Conservative 127; Mismatches 301; Indels 238; Gaps 41;

```

OY 10 LNTLNTORVLKNSK--RFTVTLVGVFLMFLALVTSNMGAKTVF-----GLVE----- 55
D 330 LSTFDTKKVTYMTSMFRSESLTSLDLSSEFTNSTETMKMLFPGANNLSLIKLOKKIKF 389
OY 56 ---SSTPAINPDSSSEYR--WGYESYVRGHPPYKQFRVAHDLRVNLEGRSRYQVCFN 110
D 390 YTVATYKLVNPPRSAPYNGTW-----TNPKTKRTILSDSLMKKDGATMGYGHWS 441
OY 111 LKRAPPLGSDSV--KKWYKKHGICSTKPEDYAMSP---RTG--DELNOKIRAVMYN 161
D 442 AVNNI--IAIDSTISGDWKAEDNFISATDQFNLIDFPLVTVGSVDTTTPGEYETIS 500
OY 162 GHPOANGIMEGLEPLAIAIRVTOEAVWYYSQNAISNDESEFRESSENLSTQSLMR 221
D 501 -----VNGLTITITVYKEMQASVVAENSTIYF--KESWK--AEDNPFYSATN---KK 545
OY 222 QALQOLIDPNLATMPKROVPDDFOL-----SIFSEDEK---GDKY 258
D 546 GKIDISSVTVTGVDVNTPDYELMYTIDGVSTKIIVYKEDKSSIEAKDSILYIGDTW 605
OY 259 NK-----GYONLLSGGLVPT--KPTPGDPRMPNPOTTSVLRKAYIGD 302
D 606 NSKDNFISANDKDNPNVDKIKVEGTVNTIKPQTNKVTYLYGQSKSEVITTVK---AD 661
OY 303 YSKLLEGATLQTLGDNVNSFOARFSSNDIGERIELSD---GYTTLTELNSPAGYSIAE 358
D 662 QSTLEADSIITYTDDKNNAKNFTATDQGNPVPDFKDIEGVGVDIT-----K 710
OY 359 PTFKVEAGKYVTTIIDQKQIENPKET-----VEPVSVEAVNDEEFSVLTONYAKFY 413
D 711 PGTNKI-----TYIYG---NLSKEVTVYKAAQATLEA-KDSALVGDWTMSKDNFIS 759
OY 414 AKNKGSSQVYVCFNADLSPDSEGGKIMTPEPT--TGEKYVTHHNGRDLFKYVPRD 472
D 760 ATDDKDG-----TVDDFKDIKVEGTVTITAGINKVYLYLGNOSKEVTVNK- 805
OY 473 TDPPTFLKHKKVLEKGYREKGAIEYSG--LTETOLRATQALAIYFTDSABLDKDL 529
D 806 ADQYTL-----EVKDSVITYTGMKAKEDNPISAT-----DKT 837
OY 530 KDYHFGEDMN-----DSTLA--VAKILVEYADSNPQLT-----DLDFLIPNNKYQSL 577
D 838 GNSVGFMDIKVEGTVTDTKAGTNKVTYTYAQSKEVITVYVADQATLE-----AKDSI 890
OY 578 I--GTQWHPEDLVDIRMEDKKEVIFVT--HNLTKRTV--TGLADRTKPFHEIEL----- 629
D 891 IYTGDKWKAED--NFISATDQAGKTIDFKNIKVEGTTI---DTTKAGNDITVYSYSGVT 943
OY 630 -KNNKOEELLSTQVTKDNLEFKDKRATINLKHGESLTLQGLPEGYSLVKE-----TDS 683
D 944 RSTELSKTITVYVAKNOVNLEAKDST-----LYGDKMIAKDNFVSATDK 988

RESULT 5
S60138
sex factor aggregation protein precursor - Lactococcus lactis
C:Species: Lactococcus lactis
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 15-Oct-1999
C:Accession: S60138; S49417
R:Godon, J.J.; Jury, K.; Shearman, C.A.; Gasson, M.J.
submitted to the EMBL Data Library, December 1993
A:Description: The Lactococcus lactis sex-factor aggregation gene cluA.
A:Reference number: S60138
A:Accession: S60138
A:Molecule type: DNA
A:Residues: 1-1243 <GOD>
A:Cross-references: EMBL:U04466; NID:9458233; PIDN:AA66196.1; PID:9458234
R:Godon, J.J.; Jury, K.; Shearman, C.A.; Gasson, M.J.
Mol. Microbiol. 12, 655-663, 1994

```



Db 1410 D-----VTENLTLDASGKLEITNLAPGDIQIETKAPAGYEIDTVPVDTITF 1461

Qy 632 NKOELLS-----OTVTDKTNLEFKOKKATINLKHGSLSLQGLPEBYSVLVETDEG 685

Db 1462 DQKLTLOVTKTNLTKTVSGKVIKAEVDPKGNV-LAEKIHIF--GI-VGDKYATKAKDKIG 1517

Qy 686 YKV---RVNSQEVANATVSKTGTSDETLAFENKKEPV-----721

Db 1518 YKLKOPNTNKGCVKFEKQK-----TFVEKKKAPIVVAPDKPVTVPKTPKVPDAKK 1571

Qy 722 -----PTGVQDKINGTALIVLAGISLGIWJHTIRIK 755

Db 1572 PTVKTSLSPTGDESPYG-----IIFTGLFASFMGFLLRKSK 1608

RESULT 7

F82884

hypothetical protein U0495 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C:Accession: F82884

R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H. submitted to GenBank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit

A:Reference number: A82870

A:Accession: F82884

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-5005 <GLA>

A:Cross-references: GB:AE002148; GB:AF222894; NID:96899495; PIDN:AAF30907.1; GSPDB:GN001

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: U0495

A:Genetic code: SGC3

Query Match 4.0%; Score 157.5; DB 2; Length 5005;  
Best Local Similarity 19.8%; Pred. No. 3.6;  
Matches 158; Conservative 100; Mismatches 285; Indels 253; Gaps 36;

Qy 5 RFPKLTMTNQRVLSKNSKRF-----VTL-----VGFLMIFALVT 42

Db 1598 RIKGTQNTISTPATISQDKRYTKCMATNLIALGDIETIVSTIKNQTKKNVSPITFELPS 1657

Qy 43 SMWGAKT---VFGLVESSTPMALNPDSSEYRWYGESYVGHYKQFVRAHDLRVNLE 99

Db 1658 GPICGYTLAPVFKTISL-----DVTYHKNKTADIKITLR 1694

Qy 100 GSRSYQVYCNLKAAPLGSDSYKWKYK-KHDSIKTFEDYAMSPRITGDELNQKLRAV 158

Db 1695 -IENIGASLFP-----KDIKFIKRRKNDGQISFIHKVTS-----1728

Qy 159 MYNCHPNANGIMEGLEPLNAIRYTOEAVWYISDNAPISNPDESFKRESNLTSTOLS 218

Db 1729 -----ANDANVEWEFKDLNRREYTERVY-----LKNKDFNQSNVSES-----1768

Qy 219 LMRQALQOLIDPNLATKMPKQVDPDFQLSIFESSEDKDKYKGYQLNLSGLVTPKPTP 278

Db 1769 -----DYIDMITEINLNKFK-----LLPTKPLGI 1793

Qy 279 GDDPMPNPQPTSVLIRKKAIGDYSKL-EGATLQLTGDNVNSFOARVSSNDIGE---334

Db 1794 IGAPIKEISNGAKVQL-KFALNDFDVLKENOTFKF---NIQENNNGANLNEISEHEG 1849

Qy 335 RIELSDG-TYTLTFLNLPAGYSTAEPITFEYKAGKVTITIDGKQIENPKKEIVE-PTISVE 392

Db 1850 KVELIDQKEFVANLN-----IKVKEKRVY--KIYFENDDVGVGYYKIN 1894

Qy 393 AYNDPEE-----FSVLTTONYA---KFYAKKKNSSOYVCFNDLKSPPSESGGK 442

Db 1895 FKNDYKRPNNVVDASVNTTQTYFTKFKALASNNLTDVANKONISINLDS-----R 1950

Qy 443 TNPDDTGGVKKYTHLAGDLFKYTVKPRDTPDTFLKHIKKYLEKGYREKGAIEYSG 502

Db 1951 VETIQGYEFKAKAYISNDNRVWYNTI-PAPTN-----IGNKNNNIALNFE-L 1995

Qy 503 TETQL---RAATOLAIYFTTSAELDKRKLKDYHFGDMNSTLAVAKILVEYADSNPP 559

Db 1996 NQNDLISRLYTLFALY---SKDVNDE-----NANNVYIKNNVPOTISTKP 2042

Qy 560 QUTDDFFIPNNKKVQSLIGTQWPEHDVJDIIMDEKKEVY-----TNHULTRK 610

Db 2043 SSTYVDLKAQNADEKKTILSLILNSDOI---FEDKNNHLIKIAKISTIDEIDAHQIINS 2098

Qy 611 TVTGLAGDRTKDFHEIELKNNKQELLSQTVKTDKTNLEFKDKATINLK-----660

Db 2099 T-----THDYDLTEKENNEMMLTKQLINL-KPNTKRYKVKWFTSKSDPTIYNGVN 2149

Qy 661 -----HGESITLQGLPEBYSVLVKEFDSEGYKKVNSQEVANATV--STGTITSD 708

Db 2150 KDNIVYPINNHNNTNIDLTLEKSPILNSVKT-----KTNFKNBSIKLMVGFNKGTSSLE 2204

Qy 709 ETLA---FENKKEPV 720

Db 2205 NKYAKLVYKDDNNNOI 2220

## RESULT 8

AH1396

peptidoglycan anchored protein (LpxTG motif) [imported] - Listeria monocytogenes (str

C:Species: Listeria monocytogenes

C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AH1396

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec

: Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kunz, M.; Kunst, F.; Karapakt, G.; Madeno, E.; Maitournam, A.;

Ok, C.; Schlueter, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AH1396

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1530 <GLA>

A:Cross-references: GB:NC\_003210; PIDN:CAD00654.1; PID:G16412064; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lmo2576

Query Match 4.0%; Score 156; DB 2; Length 1530;  
Best Local Similarity 20.6%; Pred. No. 0.69;  
Matches 160; Conservative 78; Mismatches 221; Indels 316; Gaps 38;

Qy 193 NAPISNDESEFKR--ESESNLVSTOLSIMQALQOLIDPNLATKMPKQVDDPQLSIFE 250

Db 794 NMPISPGEDYDKTGVDESN-----LEYFVKRKNELNOSYVIR-----YOTATTL 839

Qy 251 SEDK-----GDKYKNGGYQLNLSGLVPTK---PPRPDMPNPQPTSVLIRKAI 300

Db 840 TSDTETTAQIANSVTFPGDNTTKGETETKAIKVIKTTGD---GTGGETKIILNKKYDK 896

Qy 301 GDYSKLEGATLQLTGDNVNSFOARVSSNDIGERIELSDGTYTLTFLNLPAGYSI---356

Db 897 ADPSIPLGATFDLYA-NDEKVDPTQTTDKNGVIEFDLVGYDYTLKEVASAGETLTPAS 955

Qy 357 AEPTTFKE-----AGKYTTIID--GKQIEN 380

Db 956 TENIQVLEODEKVVQVWNEKMPKETGEVHLVTKRATGATTLAGAERSLYDKSGABELN 1015

Qy 381 -----PKKEIV-----EPYSVEAYNDE---EPSVLTONYAKRYVAKN---417

Db 1016 GLTIDKELTILHMLDLGSYTLAKETKAPKPEGYKLSKETWERSVESGQVDAIEIQANENEKDL 1075

Qy 418 -----NGSSQVV-----YCF-----427

Db 1076 GEAVLTIVDSEITNAKLSGAKFNLLNDSGEVIQTNLVSDENGEIRVONLEPDGVAFOGTEA 1135

```

QY 428 --NADLKS-----PPSEDC-----GKT-----443
      ||:|
Db 1136 PNNYDATTMTPTTYAGTSATMTAENNNKTKPPVDGVEVLVAKODSAGTGLGAYF 1195
QY 444 --MTPD-----FTT--GEVKYTHA-----GRDLFKYTVKKPRD 474
      ||:|
Db 1196 DLTMTAGAIVASNLTDANGELITVTLAPGKYSFKETKAPGEYELATDWEFTIADNPQE 1255
QY 475 -----PPT-FLKHKKYIEGKREKGAIESSGTEQLRAATOLAIYFT 519
      ||:|
Db 1256 KTTTAENTKLAIPAGSVKKIKODESNGVRLAG--AEFSLAENGELTQTLK--T 1309
QY 520 DSAELDKDLKDYHGFDMNDSTLAVAKILVEYADSNPPQLTDLFFIPNNKKYOSLIG 579
      ||:|
Db 1310 DEK-----GELEVNNLAPGNRI--QETKAP-----DGYO-LBS 1340
QY 580 TQWHPEDLDVITRMEDKKEVTPVTHNLTKRTV--TGLAGDRTKDFHEIEELKNNKOELL 637
      ||:|
Db 1341 TPOQFE-----IVANDTSQVTVIAENAKLEPDVAETGAVRLIKTDESETGRLSGAVFSL 1395
QY 638 SQTVKTKDNLEKDKKATINLKHGESLTLQGLPEGYSLVKTDS-EGKV-----688
      ||:|
Db 1396 DESGKVIQANLTTDE-----NGE-IFIDGLTPG-NYSLKETKAPDGEIAEOPWNRQ 1445
QY 689 -----KVNSEVANATVSKTGITSD-----ETLAFENKKEP 719
      ||:|
Db 1446 IYKGOVDVAVIIKAENSPILANGAISPGEDETDPSEIEIPVKRTDITLATEVTKLP 1500

```

RESULT 9  
AD1096  
Internalin proteins, probable peptidoglycan bound protein (LPXTG motif) homolog lmo0171

C:Species: *Listeria monocytogenes*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AD1096  
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecher, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Ertlan, K.D.; Fshl, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maltournam, A.; Mok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Ertlan, K.D.; Fshl, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AD1096  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-832 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAC98386.1; PID:g16409530; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo0171

Query Match 3.9%; Score 155; DB 2; Length 832;  
Best Local Similarity 19.9%; Pred. No. 0.3; Mismatches 244; Gaps 33;  
Matches 141; Conservative 93;

```

QY 155 LRAVWYNGHPQAN-CIMGELPLNAIRYQEAWVYSDNAPISNDSESKRESSENV 212
      ||:|
Db 176 LKSISSANNKKITGNESLVKTLPELITLEV-----LGNATITELIDENQNLVT 222
QY 213 -STOSLSMRQALKOLIDP-----LATKMPQVDDPQLSIFESE 252
      ||:|
Db 223 LSADELELKLTLKLNLSQNGLRASSISIDMGDESIVLMLPELISDISGNLDSO 282
QY 253 DKGDYKNKYONLSGLVPTKPTPGDPMPPNPQOTTSVLIRKAIG--DYSKL--- 307
      ||:|
Db 283 DILHEMLPAVKMLNDISSNELTRLPKINDEPL-----LTTINVRSKRIDRLESSKLVDVP 336
QY 308 -----EGATLOLT-----GDNNNSFGA-----RVFSSNNI-----GGR 335
      ||:|
Db 337 KLAITLADKQAVTLKTLTAAGNFTIPNNVENLAGQVTPRTIISNNGTYSQSIAMASGEL 396

```

```

QY 336 IELSDEGTYTE-LNSPA-----GYSIAEPIPEKVEAGKYVTLIDKQIE-----379
      ||:|
Db 397 SGLSKVSYTFEDEVINSPALAGKYTGIVNQPIEKA-----VPVIAKKSYSYAPVANKDEA 452
QY 380 -----NPKELVEPYSEAYNDPEEFSSVLTQNTAKFYAAKNKSGSQVYCFNA 429
      ||:|
Db 453 TFLQDIRASASEMAQITSDYS-----EVDPEATPGDYTVTLHAANKE-----F 494
QY 430 DLKSPDSDSGGCTMPPTTGEVKYTHIAGRLDFYTVKPPRDTDFLKHKKYIEKG 489
      ||:|
Db 495 DLKA-----DVTYVVAHINDIQKQ 514
QY 490 YREKGAIEY--SGLETQLRA--ATOL--AIYFTDSAEIDKDLKDYHGF--534
      ||:|
Db 515 VANNSDISPEVGETELTSEVLAKSGAVYTDLDEAKIKNEVDSEYDSSKLTGYEATIIA 574
QY 535 ---FGDMNSTLAVAKILVEYADSNPP--QLTDLFFIPNNKKY-----SLIGQ 581
      ||:|
Db 575 KSKSGASSDPKILSKVIV-----DTEKPIIQINNPETIIEKGELEGOIIDQVGITARD 629
QY 582 WHPEDL---VDITRMEDKKEVTPVTHNLTKRTV-TGLAGDRTKDFHEIEELKNNKOELL 638
      ||:|
Db 630 NYDQDLNHHMDLSKVPTSK--PGSYEVIT-----YTEDSSGNKSEIYV 670
QY 639 QTVKTKDNLEKDKKATINLKHGESLTLQGLPEGYSLVKTDS-EGKV-----691
      ||:|
Db 671 ITVKVPEARI--GKITIQYMDSENNELAESITIGGEVETELAKEI--EGYTLKEN 724
QY 692 SQ-----EVANATVS-----KITITSEDLTAFENKKEPVYPTGVDCKING 731
      ||:|
Db 725 PANSGVFEEETROTIIYVKKDIINPEFVYSNNVTPELPSNNNSVNG 774

```

## RESULT 10

AD1129  
probable peptidoglycan bound protein (LPXTG motif) lmo0435 [imported] - *Listeria mono*

C:Species: *Listeria monocytogenes*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AD1129  
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecher, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Ertlan, K.D.; Fshl, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maltournam, A.; Mok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Ertlan, K.D.; Fshl, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AD1129  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2013 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAC98514.1; PID:g16409812; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo0435

Query Match 3.9%; Score 152.5; DB 2; Length 2013;  
Best Local Similarity 19.9%; Pred. No. 1.7;  
Matches 169; Conservative 117; Mismatches 302; Indels 263; Gaps 39;

```

QY 18 VLKSNKRPFT---VTLGVFLMIFALVTSMGAKTVEGLVSSSTPNAINPOSSSEYR-- 71
      ||:|
Db 1245 LFNMSTREYTLNSVYGVGVAEIIITERSKPTTKI--LDKAVTPNATITIESSDEATEV 1302
QY 72 -----WGYESYVNGHYYKQFRANHLRVNLBESSRSYQYCFMLKKAPF- 116
      ||:|
Db 1303 TNEWTLESDGTIVYVGTGNTIR-IPNDEGTYIAKWTATDEAGNTASDEKTPDIDYVPT 1361
QY 117 --LGSDSVYK-----KWKY-----KHDGISTKFEEDA-----NSP 144
      ||:|
Db 1362 LTVYNQDASAEVNSTEANGIKRPLNAAATDTHDGNITTPYVDYSKYKMDVLGTYPTVYAT 1421
QY 145 RITGDELNOKLR-----AVMYNGHP-----ONANGIMEG--LE 175

```

```

Db      1422 DASGNKATQVNLRIVDITTSPTILITNNPLFYIENMKRLAEQELTYAAGLIGDNYDLA 1481
      176 PLNARIVTOEAVWVYSDNAPISNPDESFR---ESESUNVYSQSLMRQALQO----- 226
      1482 PGQAVQPNKOPWVFSTFIFGDIAVAKQGYOVVNLADSSG---NQALPQITITNV 1537
      227 -----LIDENLATKMPKQVDP 242
      1538 VDTGVPVIAKADNVSHVNTTTEAEFFQDARLDVTDNDPDTDLITTSNEAKVNLNRPQ 1597
      243 DPQUSIFESEDKGDKYKNGYONLSGLVPTKPPPGDPMPMPNPQPTTSVLIRRYAIGD 302
      1598 KVEVITATDTRKGNQOTFEITVVOVS---KDKPVITADPKI--SYOGKIEVEANFLSGV 1651
      303 YSKLLEGATLTGTGNVNSFQARVPSNDIGERIELSD-GYTTLELMSPGAY-SIAREI 360
      1652 HAEV---TDELDDG-----VAKTSDFAEKVDENKVGTYVTF-LNAKDEYGMTAEV 1698
      361 TPKV---EAGKVVYIIDGKOIENPNKEIVEPYSEAYANDPEEFSVLTTQNAKPYAKN 416
      1699 KVSYSIFPKKIAPTFENNAKNTIEAVN---ELPSLESTFKIE-----AKDYLSGN 1744
      417 KNGSSQVYVYCFENADIKSPDSE-----DGGK-----TMDPDTTG-EVKYTHIAG 460
      1745 K---LKVTYTPEQITKGNVPEYSIKVTTKDGSNIAETVTLTIKDTGTSPIKMT----- 1797
      461 RDLFKRYTYKPRDPTDPTLKHKKYIEKGYREKGAIEYSGLTETQLRAATQALYY-FT 519
      1798 KSTKLEVOAKEPMNITFEFGIKADVIDVDVTKNKIKVDS---EVNLNKKVGTYPTFYVYT 1854
      520 DSAELDKDLKLDYHFGGDMNSTLAVAKILVEYAODSNPOLTDLDFIPNNKRYQSLIG 579
      1855 DALGNESKSLSTQVIVTSSBELTIDKLEISTYPIGKI---ISDKQILODICTKYTNISYG 1911
      580 TQWMEDELVDIIRMEDEKKEVIPVTNHLT-----LKRKYT-----GLAGDRTKDFH 624
      1912 T-----VKVTNLSKIYDMNKAQYKVTATNTSSGVAEKT----- 1948
      625 FEIELKNNKQELISQVATKDTNLEFK-----DCKATINLKHGSLTLOGLPEYSILV 678
      1949 ILLVYKNSDSSFIAPNSKDDKNNKPAKNIPTGTDLNTELJVMGMLLLVG---GWMFLR 2005
      679 KETDESGYKVK 689
      2006 RKT-----KVK 2011
      Db

```

RESULT 11  
 S70172  
 C:Species: Clostridium difficile  
 C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 15-Oct-1999  
 C:Accession: S70172; S44271  
 R: von Elchei-Streiber, C.; Meyer zu Heringdorf, D.; Habermann, E.; Sartling, S.  
 Mol. Microbiol. 17, 313-321, 1995  
 A:Title: Cloning in on the toxic domain through analysis of a variant Clostridium difficile  
 A:Reference number: S70172; MUID:96079281; PMID:7494480  
 A:Accession: S70172  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-2367 <YON>  
 A:Cross-references: EMBL:223277; NID:g761713; PIDN:CAA80815.1; PID:g761714  
 A:Experimental source: Isolate 1470  
 R: Sartling, S.; von Elchei-Streiber, C.  
 submitted to the EMBL Data Library, July 1993  
 A:Description: Cloning and sequencing of an Clostridium difficile toxin B mutant.  
 A:Reference number: S44271  
 A:Accession: S44271  
 A:Molecule type: DNA  
 A:Residues: 1-1323,'N',1325-2367 <SAR>  
 A:Cross-references: EMBL:223277  
 C:Superfamily: cpl repeat homology

C:Keywords: cytotoxin

Query Match 3.8%; Score 151; DB 2; Length 2367;  
 Best Local Similarity 19.9%; Pred. No. 2.7;  
 Matches 171; Conservative 94; Mismatches 292; Indels 304; Gaps 42;

```

      5 REPNNLNTLNTQRLV-----SKNSKRYTTLGVFLMIFALYTSWGAQTVGLVE--SS 57
      1568 KPMNRKGSNTSDSLMSFLESNMKISFVNFQSNKIFLDANFIISGTSISGQFEFID 1627
      58 TPNALNP-----DSSSEYKRY-GYESTYKGPYTKQFVAADLRVNLGSSSYOVCN 110
      1628 ENNNIOPYFIKENTLETNTVLYGNKQNNIVAPNY-----DLDDSDISSTVIN 1676
      111 LKAPPLGSDSVKMKYKKHGDISKRFEDYASPRTQDELQKLRVANNHQPQ-----N 166
      1677 FSOKLYLGIDSCVNK-----VVISPNLYTDEINITPYETNNYPEYIVLD 1722
      167 ANIGMEGLEPLAIRVTOEAVWVYSDNAPISNPDESFKRESNLSVTSQSLMRQALQO 226
      1723 ANYINKEIN-VAINDSLIRYV-----SNDGND-----ILMSTSEKVKQVKT 1767
      227 LI-----DPLATKMPKQVPPDFQLSIFESEDKD-----KYNKYQNLISG-- 268
      1768 FVNVFEKDKTLANKLS-----FNSDKQDVPVSEILISFTPSY---YEDGLIGYD 1813
      269 -GLVPTTKPPTPDPPMPNPQPTTSVLIRKVAIGDYSKLBEATLQLGDNVNSFQARV 327
      1814 LGLV-----SLYNEKFFYNNFGMVSG--LIYINDSLYKRPV- 1850
      328 SINDIGERIELSDGYTTLELMSPGAYSIAEPI-----TFKVAGKRY 370
      1851 -NNLTGVTVDQDKYFNPINGAA-SIGETIIDDKNYFNQSVLQGVSTEDGFKY 1908
      371 -----TI-IDGKOIENPNKEIVEP---YSVEAYANDEEFSVLTTQNY-----AKFY 413
      1909 FAPANTLDLENLEGEAIDFTGKLIDENIYFEDNRYGAVEKMLEGEMHYFSPETGAKFK 1968
      414 AKNKGSSQVYVYCFNAD--LKSPPDSEOGKTMPTDFTTGKRYTHIAR----- 461
      1969 GLNQIGDCK--YFNSDGYMQGFVSIQNNKHYFSDSGVAKGYEIDKHHYFAENGEM 2026
      462 -----DLFKYTYKPRDPTDPTLKHKKYIEKGYREKGAIEYSGLTETQLRAATQ 513
      2027 QIGVFNTEDGFKY-----FAHNHEDL-----GNREGELISYSGLNNK----- 2066
      514 AIYFTDSAE-----LDKDLKLDYHGFQDMN-----DSTLAAVAKILV 550
      2067 -IYFDDSFYAVVGMKLEDGSKYYFEDDETAAYIGLSLINDGQYFYNDGIMQGEVYTI 2125
      551 E-----YAODSNPQLTDL-----FTIPANN----- 572
      2126 NDKVYFSDSGILIESGYONIDNRYFIDNGIVQIGVFETSDGYRYFAFANTVNDNIYQ 2185
      573 --KYOSL-----IGTQWMEDELVDIIRMEDEKKEVIPVTNHLTRKTYTGL 615
      2186 AVEYGLVAVGVDVYFGETYITFEGW-----IYDHENESDKYFFPET-----KKAACKI 2236
      616 AGDRTKDFHE-----IELKNNKQELISQVATKDTNLEFKDGAKATMLK--HGE 663
      2237 NLIDDIKYYFDEKGIIMRTGLISFENNYYF-----NENGEIGFYINIEDKKHFEYGE 2288
      664 SLTLQ-----GLPEGSYLYKE 680
      2289 DGMQIGVFNTPDGFKRYFAHQ 2309
      Db

```

RESULT 12  
 T28679  
 C:Species: Staphylococcus aureus  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-May-2000  
 C:Accession: T28679







Db 466 VLVRKNNNSVATSKTYKVDANKLTONGKVDVKEPDSLNSKEYIIEMEKLVVDGKEVYTD 525  
Oy 448 ---FTTGFKV-YTHIAGRDLEFKTYKPRDT-----DPDFIKHIKKV--IEKGRKRG 494  
Db 526 NWYTIASLTAKKPSMAGIDLSYSTKEGFTAVPILIDPNSITNIRYVAILEEDYKANG 585  
Oy 495 QAIE---YSGLETOLRAATOLA-----IYFTDS----- 521  
Db 586 QMAEYAVSYVNASOKKAAVKGRTVDMSDGNVFAVAISGNNNSDFTLATPVSNAVYV 645  
Oy 522 -----AELDKDLK-DYHGF-----GDMNDSTLAVAKTLVEYADSDNPQ 560  
Db 646 GLSKPVEFSLEKAEODKLTINYEVPADNTLLYDNLTHPLTKLYKSDAOGMYGSGEVA 705  
Oy 561 LTLDLFFIPNNKKQSLI-----GLOHREDVDIIMMEKK 597  
Db 706 TVDL-----RNKNEITNLLEFGLESEYVVVVMATASYLDGAGIMVOLLGQSGVFQTK 761  
Oy 598 EVIIVTNHLTLRKVTGLAGDRTKDFHEIELKNNKQELISQYKT-DKTN-----LEF 650  
Db 762 EIAVNAKFTLESYTNNA-----EVNIKLEPAKLNLSANLKYKKNNTLVKTYPL 814  
Oy 651 KDGRATINLKHGESLTLQGLPEGYSLVKEPDESGYKVKVNSQEVANATVSKT---GIT 706  
Db 815 NDDFNKLTMTEDGYTYLFEELAIKKEYLKVED--GLDSGMNOVPVEGQLVFKTKASPLT 872  
Oy 707 SDETLAFENKKEPVVPTGVODKINGYLALEYIAGI 741  
Db 873 DKVLLDYQPNK-----LKVGG-----LAGI 892

## RESULT 14

hypothetical protein U0558 [imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: G82875  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.  
Submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit  
A:Reference number: A82870  
A:Accession: G82875  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1883 <GLA>  
A:Cross-references: GB:AE002154; GB:AF222894; NID:g6899557; PIDN:AAF0971.1; GSPDB:GN001  
A:Experimental source: Bioovar 3; Bioovar 1  
C:Genetics:  
A:Gene: U0558  
A:Genetic code: SGC3

Query Match 3.8%; Score 148.5; DB 2; Length 1883;  
Best Local Similarity 19.1%; Fred. No. 2.7;  
Matches 155; Conservative 102; Mismatches 276; Indels 277; Gaps 34;  
Oy 103 SYOYCFELKKRA--PLGSDSSVKKMYKKHGDISTFEDVAMS----- 143  
Db 951 NFDISMFTLKYTFDHPMTSMKRMARITDAMKTYDFDNISKLKIDEKANASSVRY 1010  
Oy 144 ---PRITGDELNOK-----LRVYNGHPONANGIMGELPLNAIRVTOEAVWYS 191  
Db 1011 KKVLSQILIGEDITNKTFOFYALLTLQYNASKIWDNKSKEVSKLDELII--EIKKYY- 1067  
Oy 192 DNAPISNDESKFRSESNLYSTQLS-----MRALQOLIDPNLAT 234  
Db 1068 -----DKQFDRGVANNVVAQTLLTFNFDAGIEGNGKYYFKVEKEEELYNILMT 1119  
Oy 235 KMPKQVPD-----DQLSIFSESDGDKYKNGY 262  
Db 1120 KTKESVSEIIGIRPSYLTQNKITSYEOLEFNKSMILNFGITIKLVADGDKGNKHLKL 1179  
Oy 263 QNLLSGGLVPTKPTPGDPMPMPNQPTTSVLIRKYYAIG-DYSKILEGATQIOLGDNVNS 321

Db 1180 DELTEKEIYPLMIITPAE-----AQAINFDYKNSNGSFPTPLEN-NG 1222  
Oy 322 FQARFSSNDIGERIEISDGYTTLTFLNSPAGISIAEPIITFEVGAQYTIID---GKOI 378  
Db 1223 SKFYTIKRDIPSLIEF-----MSVDPK-YTIVENALSSSH 1258  
Oy 379 ENPNK-----EIVEPYSV-----EAYND--FEFESVLT-----ONVAKF 411  
Db 1259 ENVRKWDYELKERDVKFNEVIAKKQEAKNLLEBFQKOLLTGLDGFLESNDIRKF 1318  
Oy 412 YVAKKNGSSQVY--CFNADIKSPDSEGGKMTPTPOTGEVYTH-IGARDLKYTV 468  
Db 1319 YKSKDFNPSLDKYQVDFDKIGLYGFSFGKRYERPSPYGEPLDTYVFNAGNPQVWKA 1378  
Oy 469 KPRDTPD---POTFLKHKKVIEKGRK-----GGAIEYS-GLT----- 503  
Db 1379 KPNQKNVPTVSIKIGQLEIERKRENTLNFGLQDYARFETTYTDI PGSSKDIYWL 1438  
Oy 504 -----ETQLRAATOLAIYPTDSAEIDKDKIDYHGFGDMNDSTLAVAKILV-EY 552  
Db 1439 GSFEGISOSSSELPNADGIVSSDYTLKKRKQKQELNIEFG-----DYFNIAEVLTRDY 1493  
Oy 553 AQ-----DSNPQOLTDL-DEFIPNNKKYSLIGTQWHPEDLVLIIRMEKKEYIP 601  
Db 1494 VQTVFIPSONELDNLPVYISGLSDPNTGNEYVFGSDNTKQNN-----ERLIP 1540  
Oy 602 VTHNLTUKTYTGLAGDRTKDFH-----FEIELKNNK-----QELISQYK 642  
Db 1541 VNNFLSVSNTJIALNTLFAITNNKELQNVLANQALVYQKNOOLFDDYTLTKPOELLESNNK 1600  
Oy 643 TDKTNLEFKDGRATINLKHGESLTLQGLPEGYSLVKEPDESGYKVKVNSQEVANATVSK 702  
Db 1601 VKESFLSANPLAL--TTHDSVDVLANL-NSISFLTRLSPK-----SNSYIGK 1645  
Oy 703 TGIITSD-----ETLAFENKKEPV 720  
Db 1646 TRLTNNGFEDRWLRKIIDWEIYDNRBPI 1675

## RESULT 15

cell wall-associated protein precursor wapa (B. subtilis) homolog lin0454 [imported]  
A:1489  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AF1489  
R:Glaser, P.; Fraigneul, L.; Buchrieser, C.; Amend, A.; Baguerio, F.; Berche, P.; Bloec  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fslini,  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;  
Ok, C.; Schlueter, T.; Simoes, N.; Tietzer, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AF1489  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2167 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CAQ95686.1; PID:g16412895; GSPDB:GN00178  
A:Experimental source: strain Cl1p11262  
C:Genetics:  
A:Gene: lin0454

Query Match 3.7%; Score 147.5; DB 2; Length 2167;  
Best Local Similarity 19.0%; Fred. No. 3.8;  
Matches 179; Conservative 111; Mismatches 319; Indels 331; Gaps 46;  
Oy 47 AKTVFGL-----VESSTPNAINPDSSSEYRVGYEYRGHPYKOF-RVAAD--- 93  
Db 562 AGSTGLNLNPLRYLFRVQAVFPGGSPSNDLVFAFMLEK--PQAPVAKAVASNLNKHKG 619  
Oy 94 -LRVNLSS--RSYQVCFNLK--KAFPLGSDSSVKKMYKKHDI-STKEF---DYAM 142  
Db 620 YVELNMEKSPMADGYKVLVFNKGAYEEYDGAET---KWTQNKIGIMPTKEIEAGKYAL 676

QY 143 -----SPRITSDENOK-----LRAVMYNGHPONANGIMEGLEP----- 176  
Db 677 HHDGKAELAKDPSPYTNSGVNKTERTNYWFRVIAQKAGNNMTSVQS--EPATPSIPE 734  
QY 177 -----LNAIR---VTOFAVWYISDNAPISNPDESFKRESE-SNL 211  
Db 735 ALNKGQMDVWYTSVPYRGGEVNAATNGNPLFHEHDFNLGCRPSINVRNRTFNSODDATGI 794  
QY 212 VSTQSLSMRQALKOLIDPNIATKMPKQVPPDDPOLSFESSEDGDKY----- 258  
Db 795 FGKGTSTLEEKLYEENGNIY-----WVESDKIHRTK--KGDYEAPPGIYSEITKN 847  
QY 259 -----NKGYONLSGGLVPTKPTPGDPPMPNOQOTTSVLRKVAIGDYSKILE- 308  
Db 848 ADGYLKEEDKSETRFLVDGRKSEKDTKGN-----LTVEYTDGKLTSLRDA 895  
QY 309 -GATLOLT--GDNV-----NSFOARVFSNDIGE--RIELSDGYT--- 344  
Db 896 SGRVTYLTLYEGELVKELVGEDRRIKYNDKQELISSYARGLYRYGYTDGILTSYD 955  
QY 345 -----LTENSPAG-----YSIAEPITFKVEAGKYTTIIDGKQIE 379  
Db 956 PKHTEKPYETTPAYEEKLTETIDPYGKKTTLSTYDKAEQOTTLTNEKKKTTYSYNDAG 1015  
QY 380 NPKKEIVE-----PYVEAYNDEEFSVL--TTONYA-----KPYAKNKGSSQVY 425  
Db 1016 NPKKEIVDADGLKLTYYTESNNLVKENVKGOEETAYADAGNITKATDAYGTESYTY 1075  
QY 426 CFNAADLKSPDSEDGCKTMTPDFTTGCVKTYTHIAGRDLFKYTYVPRD-----TDPDTFLK 480  
Db 1076 NDNNDYTSSYDTGKRTTYTD-----GADAVSETLATESQVSSVTOYDAYGN 1123  
QY 481 HIK-----KVIENGYREKGAIEYSGLTETOLRAA-----TOLA--IYFFDSAEI 524  
Db 1124 PIRGSGELSSGGLNLSNGFEGAGVSNMTLIQSDAKGSMTFDNTOSAPGALGSGSVKL 1183  
QY 525 DKDKLDYHGFQDMN-----DSTLAVA-----KILVEYQDSNPOLTLD 565  
Db 1184 TSEANSTYKGYSSVTOVRDEPETTYTFSAMIKTSGMTNADALLIGRLQDANKAKDYTDAG 1243  
QY 566 F-----FIPNNKYOSLI-----GTQWHPEDLVDT 591  
Db 1244 VMOSNRATSIKKNGDWVKROLTFKTSKNTROVLLYLDNEQAPAPHKGKITWY--DNVOFE 1301  
QY 592 RMEDKKEVLPVTHNLTKRTVTLGADRTKDFHEIELEKNNKQELLSQTVKTKTNLEFK 651  
Db 1302 KGSVASSYNPVVNN-----SEENHNGTLLPTGMWRTGNTAL--T 1337  
QY 652 DGRATINLKH--GESLTL--QGLPEGYSYLKETSDEGYKVKVNSOEVANATVSKTGITS 707  
Db 1338 QAKVVDNOSHSGDSAYFEERKATSEAYTHIVOD-----VPYNOKAKALITIS--ALSK 1388  
QY 708 DETLAFENNKPEVPTGVOKINGYLALIYAGISLGIWG 747  
Db 1389 SE-----DAKANGSVATM--SNDYSVWG 1409

Search completed: August 19, 2003, 15:22:29  
Job time : 48 secs